A Robust Optimization Approach for Dynamic Traffic Signal Control with Emission Constraints

Ke Han[a*] Terry L. Friesz[b†] Hongcheng Liu[b‡] Tao Yao[b§]

[a] *Department of Mathematics*
Pennsylvania State University, PA 16802, USA
[b] *Department of Industrial and Manufacturing Engineering*
Pennsylvania State University, PA 16802, USA

Abstract

We consider an adaptive signal control problem on signalized network using Lighthill-Whitham-Richards (LWR) model (Lighthill and Whitham, 1955; Richards, 1956), with traffic-derived emission side constraints. We seek to tackle this problem using a *mixed integer mathematical programming* approach. Such a problem class, which we call LWR-Emission (LWR-E), has been analyzed before to some extent. As pointed out by Bertsimas et al. (2011), the mere fact of having integer variables is not the most significant challenge to computing solutions of an LWR-based mathematical programming model; rather, it is the presence of the nonconvex emission-related constraints that render the program computationally expensive.

As we will show in this paper, it is computationally practical to solve the LWR-E problem when it is expressed as a mixed integer program. In particular, we propose a *mixed integer linear program* (MILP) for the LWR-E which can be efficiently solved with commercial software. The proposed MILP is meant to explicit capture vehicle spillback, avoid traffic holding, and reformulate emission-related side constraints as convex and tractable forms. The primary methodologies employed by this paper are: (1) a link-based kinematic wave model (Han et al., 2012); (2) an empirical and statistical observation between aggregated emission rate and certain macroscopic traffic quantity; and (3) robust optimization techniques.

1 Introduction

Traffic signal is an essential element to the management of the transportation network. For the past several decades, signal control strategies have evolved from ones developed based on historical information, often referred to as the fixed timing plan, to the generation of control strategies in which the control system is fully responsive. In the latter case, the cycle lengths and splits of the signal are determined based on real-time information. Representatives of such signal-control systems are OPAC (Gartner, 1983), RHODES (Mirchandani and Head, 2000), SCAT (Sims and Dobinson, 1980) and SCOOT (Hunt et al., 1982).

*Corresponding author, e-mail: kxh323@psu.edu;
†e-mail: tfriesz@psu.edu
‡e-mail: hql5143@psu.edu
§e-mail: tyy1@engr.psu.edu

1.1 Mathematical programming approach to traffic signal control

The performance of a traffic signal control system depends on the optimization procedure embedded therein. We distinguish between two optimization procedures: 1) heuristic approach, such as those developed with feedback control, genetic algorithms and fuzzy logic; and 2) exact approach, such as those arising from mathematical control theory and mathematical programming. Among these exact approaches, the *mixed integer programs* (MIPs) are of particular interest and has been used extensively in the signal control literature. Improta and Cantarella (1984) formulated the traffic signal control problem for a single road junction as a mixed binary integer program. Lo (1999a) and Lo (1999b) employed the *cell transmission model* (CTM) (Daganzo, 1994, 1995) and casted a signal control problem as mixed integer linear program. In these papers, the author addressed time-varying traffic patterns and dynamic timing plan. In Lin and Wang (2004), the same formulation based on CTM was applied to capture more realistic features of signalized junctions such as the total number of vehicle stops and signal preemption in the presence of emergency vehicles. One subtle issue associated with CTM-based mathematical programs is the phenomena known as traffic holding, which stem from the linear relaxation of the nonlinear dynamic. Such an action induces the unintended holding of vehicles, i.e., a vehicle is held at a cell even though there is capacity available downstream for the vehicle to advance. The traffic holding can be avoided by introducing additional binary variables, see Lo (1999c). However, this approach ends up with a significant amount of binary variables and yields the program computationally demanding. An alternative way to treat holding problem is to manipulate the objective function such that the optimization mechanism enforces the full utilization of available capacities in the network. This approach however, strongly depends on specific structure of the problem and the underlying optimization procedure. Specific discussion on traffic holding can be found in Shen et al. (2007)

This paper is concerned with controlling signalized junctions where the dynamics of vehicular flows are governed by the network extension of the Lighthill-Whitham-Richards model. In particular, we employ the *link-based kinematic wave model* (LKWM) proposed by Han et al. (2012). This model describes network dynamics with variables associated to the entrance and exit of each link. It employs a Newell-type variational argument (Newell, 1993; Daganzo, 2005) to capture shock waves and vehicle spillback. Analytical properties of this model pertaining to solution existence, uniqueness and well-posedness are provided in Han et al. (2012). A discrete-time version of the LKWM, known as the *link transmission model*, was discussed in Yperman et al. (2005). In contrast to the cell-based math programming approaches where the variables of interest correspond to each cell and each time interval, the model proposed in this paper is link-based, i.e. the variables are associated with each link and each time interval. The resulting MILP thus substantially reduces the number of (binary) variables and hence the computational effort. In addition, the link-based approach prevents vehicle holding within a link without using binary variables. The network model employed in this paper captures key phenomena of vehicular flow at junctions such as the formation, propagation and dissipation of physical queues, spillback and vehicle turning. It also considers important features of signal control such as dynamic timing plan and time-varying flow patterns.

1.2 Lighthill-Whitham-Richards model

Following the classical model introduced by Lighthill and Whitham (1955) and Richards (1956), we model the traffic dynamics on a link with the following first order *partial dif-*

ferential equation (PDE), which describes the spatial-temporal evolution of density and flow

$$\frac{\partial}{\partial t}\rho(t,\, x) + \frac{\partial}{\partial x} f\big(\rho(t,\, x)\big) \;=\; 0 \tag{1.1}$$

where $\rho(t,\, x) : [0,\, +\infty) \times [a,\, b] \to [0,\, \rho_j]$ is average vehicle density, $f(\rho) : [0,\, \rho^{jam}] \to [0,\, C]$ is average flow. ρ^{jam} is jam density, C is flow capacity. The function $f(\cdot)$ articulates a density-flow relation and is commonly referred to as the fundamental diagram.

Classical mathematical results on the first-order hyperbolic equations of the form (1.1) can be found in Bressan (2000). For a detailed discussion of numerical schemes for conservation laws, we refer the reader to Godunov (1959) and LeVeque (1992). A well-known discrete version of the LWR model, the Cell Transmission Model (CTM), was introduced by Daganzo (1994, 1995). PDE-based models have been studied extensively also in the context of vehicular networks, with a list of selected references including Bretti et al. (2006); Coclite et al (2005); Daganzo (1995); Herty and Klar (2003); Holden and Risebro (1995); Jin (2010); Jin and Zhang (2003); Lebacque and Khoshyaran (1999, 2002).

Let us introduce function $N(\cdot,\, \cdot) : [0,\, +\infty) \times [a,\, b] \to \mathbb{R}$, such that

$$\frac{\partial}{\partial x} N(t,\, x) \;=\; -\,\rho(t,\, x), \qquad \frac{\partial}{\partial t} N(t,\, x) \;=\; f\big(\rho(t,\, x)\big) \tag{1.2}$$

Function $N(t,\, x)$ is sometimes referred to as the Moskowitz function or the Newell-curves. It has been studied extensively, for example by Claudel and Bayen (2010); Daganzo (2005); Moskowitz (1965); Newell (1993). A well-known property of $N(\cdot,\, \cdot)$ is that it satisfies the following Hamilton-Jacobi equation

$$\frac{\partial}{\partial t} N(t,\, x) - f\left(-\frac{\partial}{\partial x} N(t,\, x)\right) \;=\; 0 \qquad (t,\, x) \in [0,\, +\infty) \times [a,\, b] \tag{1.3}$$

1.3 Traffic operation with environmental constraints

Unlike many existing literature on signal control which primarily focus on minimizing congestion level and travel time, this paper address environmental sustainability in traffic operation and management by consider vehicle emission-related side constraints. However, as pointed out by a survey paper Szeto et al. (2012) and the literature therein, environmental considerations result in nonlinear and nonconvex feasible sets and objective functions in the mathematical programming formulation, adding computational challenges to the problem. As a result, heuristic methods (Ferrari, 1995) have been developed for these problems. Classical methods such as the inner penalty technique Yang and Bell (1997) and augmented Lagrangian multiplier technique Yang et al. (2010) have also been used.

This paper presents a novel approach to circumvent the aforementioned computational challenge by re-formulating the emission side constraints as linear constraints, using empirical observations and robust optimization. Our analysis relies on an observed relationship between aggregated link emission rates and the link occupancy, when certain velocity-based emission models are employed. Such empirical data are obtained through extensive numerical simulation. Detailed description of the simulation and synthetic data is presented in Section 5.

Despite the strong correlation between the aggregated emission rate and certain macroscopic traffic quantities (e.g. link occupancy), there are non-negligible errors associated with such a relationship. These errors are treated as modeling uncertainty and can be naturally handled by a technique called *robust optimization* (RO), which is explained in the next introductory section.

1.4 Robust optimization

Uncertainty is usually not negligible in modeling. Errors and perturbations to a deterministic model can render an optimal solution in the ideal case suboptimal in implementation. A natural approach to capture uncertainty is by assuming that uncertain parameters follow certain probability distributions and by employing the notions and methodologies in stochastic programming. However, such an approach has arguably two limitations: 1) Exact knowledge of distributions is often difficult to acquire, and 2) stochastic programming is recognized as highly intractable to solve even with linear objective function and linear constraint functions. In view of these challenges, we propose to handle uncertainty in the perspective of robust optimization.

A robust optimization is a distribution-free uncertainty set approach that seeks to minimize the worst-case cost and/or to remain feasible in the worst scenario. Compared to stochastic programming, robust optimization makes no assumption on the underlying distribution of uncertain parameters. Moreover, it has been shown to work as a powerful approximation to stochastic programming and even probabilistic models with significantly reduced computational cost (Ben-Tal and Nemirovski, 1998, 1999, 2000; Bertsimas et al., 2011a,b; Bandi and Bertismas, 2012; Rikun, 2011). Although solutions to robust optimization problems can be relatively conservative, the conservatism is adjustable with the flexibility of choosing uncertainty sets (Bertsimas and Sim, 2004). A comprehensive review of robust optimization is provided by Bertsimas et al. (2011a).

In this paper, we will invoke the robust optimization approach to capture the errors arising from statistical learning of the emission model. The goal of this approach is to maximize the total throughput of the signalized network while keeping the vehicle emission below a desired level even in the worst case.

1.5 Oganization

The rest of this paper is organized as follows. In section 2, we present the link-based kinematic wave mode (Han et al., 2012). Section 3 presents the basic mixed integer linear programming formulation for signalized network, without any emission constraints. Section 4 extends the MILP obtained from Section 3 to incorporate emission side constraints, using statistical learning and robust optimization in a general setting. In Section 5, we present details of a numerical simulation that gives rise to an empirical observation of relationship between certain aggregated quantities. Such an observation provides support to our RO-based approach. Section 6 presents a numerical example, which demonstrates and evaluates the proposed methodology.

2 Link-based Kinematic Wave Model

In this section, we present a kinematic wave model on networks, with a triangular fundamental diagram for each link. Unlike the cell-based models Daganzo (1994, 1995), the proposed model does not require modeling or computation in the interior of the link. For this reason, we call it the link-based kinematic wave model.

2.1 State variables of the system

Consider the link represented by an interval $[a, b]$, with $b - a = L > 0$. In the derivation of the LKWM, we select flow $q(t, x)$ and regime $r(t, x)$ as the state variables for the link, instead of density. It is obvious that a single value of q corresponds to two traffic states:

1) the free flow phase ($r = 0$); and 2) the congested phase ($r = 1$). Therefore, the pair $(q(t, x), r(t, x)) \in [0, C] \times \{0, 1\}$ determines a unique density value. This simple observation gives rise to the following map

$$\psi(\cdot) : [0, C] \times \{0, 1\} \ \rightarrow \ [0, \rho^{jam}], \qquad (q, r) \mapsto \rho \tag{2.4}$$

2.2 Riemann problem at a junction with one incoming link

Extension of the kinematic wave model to a network turns out to be subtle; the issues associated therein are 1) a proper definition of a weak entropy solution at a junction of arbitrary topology; 2) uniqueness and well-posedness of the entropy solution. The reader is referred to Han et al. (2012); Garavello and Piccoli (2006); Jin (2010); Jin and Zhang (2003) for some specific discussion. A junction model can be analyzed by considering a Riemann problem, which is an initial value problem with constant datum on each incoming and outgoing link. Due to space limitation, instead of a comprehensive discussion on various types of Riemann problems, we focus on the Riemann problem for a particular junction, that is, the one with one incoming link and $n \geq 2$ outgoing links. This is because we assume that during one signal phase, cars from only one incoming link can enter the junction.

In order to model vehicle turning, we fix a traffic distribution matrix

$$A \ = \ \begin{pmatrix} \alpha_{1,2} & \alpha_{1,3} & \dots & \alpha_{1,n+1} \end{pmatrix}$$

where $0 \leq \alpha_{1,i} \leq 1$, $i = 2, \dots, n+1$, $\sum_{i=2}^{n+1} \alpha_{1,i} = 1$. The coefficients $\alpha_{1,i}$ determines how the traffic from the incoming link I_1 distributes in percentages to the outgoing link I_i. For simplicity, we assume A is time-independent. Note that there is no substantial difficulty with transforming our modeling framework to deal with time-varying distribution matrices; such extension, however, requires additional information on route choices, which is beyond the scope of this paper.

The next theorem characterizes the solution to the Riemann problem at junction with one incoming link.

Theorem 2.1. *Consider a junction with one incoming link I_1 and $n \geq 2$ outgoing links $I_2, \dots, I_{n+1}$. For every initial data $y_{1,0}, \dots, y_{n+1,0} \in [0, C] \times \{0, 1\}$, there exists a unique $n+1$-tuple*

$$\hat{y}_1, \ \dots, \ \hat{y}_{n+1} \in [0, C] \times \{0, 1\}$$

where $\hat{y}_i = (\hat{q}_i, \hat{r}_i)$, such that the solutions to the initial-boundary value problems at the junction

$$\begin{cases} \frac{\partial}{\partial t}\rho(t, x) + \frac{\partial}{\partial x} f_1\big(\rho(t, x)\big) \ = \ 0 \\ \rho(0, x) \ = \ \psi\big(y_{1,0}\big) \\ \rho(t, b_1) \ = \ \psi\big(\hat{y}_1\big) \end{cases} \qquad \begin{cases} \frac{\partial}{\partial t}\rho(t, x) + \frac{\partial}{\partial x} f_i\big(\rho(t, x)\big) \ = \ 0 \\ \rho(0, x) \ = \ \psi\big(y_{i,0}\big) \\ \rho(t, a_i) \ = \ \psi\big(\hat{y}_i\big) \end{cases} \qquad j \ = \ 2, \dots, n{+}1$$

is the admissible weak solution to the junction problem in the sense defined in Coclite et al (2005). In addition, we have the following characterization: the boundary states $(\hat{q}_i, \hat{r}_i)$, $i = 1, \dots, n+1$ are given by

$$\hat{q}_1 \ = \ \min\left\{ q_1^{max}, \ \frac{q_2^{max}}{\alpha_{1,2}}, \ \frac{q_3^{max}}{\alpha_{1,3}}, \ \dots, \ \frac{q_{n+1}^{max}}{\alpha_{1,n+1}} \right\} \tag{2.5}$$

$$\hat{r}_1 \ = \ \begin{cases} 1, & \text{if} \quad r_{1,0} \ = \ 1 \\ 0, & \text{if} \quad r_{1,0} \ = \ 0, \ \hat{q}_1 \ = \ q_{1,0} \\ 1, & \text{if} \quad r_{1,0} \ = \ 0, \ \hat{q}_1 \ < \ q_{1,0} \end{cases} \tag{2.6}$$

$$\hat{q}_i \;=\; \alpha_{1,i}\,\hat{q}_1, \qquad i \;=\; 2,\,\ldots,\,n+1 \tag{2.7}$$

$$\hat{r}_i \;=\; \begin{cases} 0, & \text{if} \quad r_{i,0} \;=\; 0 \\ 1, & \text{if} \quad r_{i,0} \;=\; 1,\; \hat{q}_i \;=\; q_{i,0} \\ 0, & \text{if} \quad r_{i,0} \;=\; 1,\; \hat{q}_i \;<\; q_{i,0} \end{cases} \tag{2.8}$$

where

$$q_i^{max} \;\doteq\; \begin{cases} q_{i,0} + r_{i,0}(C_i - q_{i,0}), & i \;=\; 1 \\ C_i + r_{i,0}(q_{i,0} - C_i), & i \;=\; 2,\,\ldots,\,n+1 \end{cases} \tag{2.9}$$

Remark 2.2. *The quantity q_i^{max} is the maximum flux an incoming (outgoing) link can send (receive) – a quantity identified as demand (supply) by Lebacque and Khoshyaran (1999, 2002).*

The aforementioned map $(y_{i,0})_{i=1,\ldots,n+1} \mapsto (\hat{y}_i)_{i=1,\ldots,n+1}$ is commonly referred to as the Riemann solver, see Garavello and Piccoli (2006) for a formal discussion. Theorem 2.1 describes the Riemann solver using the new state variables q and r. The verification of theorem 2.1 is straightforward. For junctions with arbitrary topology, the Riemann Solvers are not available in closed-form. Yet, in the case of one incoming link, we are able to express the Riemann solver explicitly.

In the next subsection, we analyze shock formation and propagation within one single link. The location of the shock wave is crucial as it determines the regime variable r at the two boundaries of the link.

2.3 Shock formation and propagation within the link

We focus on solutions generated by assuming an initially empty network, i.e. $y_{i,0} = (0,\,0)$. The key to our analysis is the location of a so-called *separating shock*, which divides each link into two zones: free flow zone ($r = 0$), and congested zone ($r = 1$). We begin with the fact that if the network is initially empty, then there can be at most one separating shock on each link

Lemma 2.3. *For every link I_i and any solution $y_i(t,\,x) = \big(q_i(t,\,x),\,r_i(t,\,x)\big)$ with $y_i(0,\,x) = (0,\,0)$, the following statement holds:*

1. *For every $t \geq 0$, there exists at most one $x_i^*(t) \in (a_i,\,b_i)$ such that $r_i(t,\,x^*(t)-) < r_i(t,\,x^*(t)+)$*
2. *For all $x \in [a_i,\,b_i]$,*

$$\begin{aligned} r_i(t,\,x) \;&=\; 0, \qquad \text{if} \;\; x \;<\; x_i^*(t) \\ r_i(t,\,x) \;&=\; 1, \qquad \text{if} \;\; x \;>\; x_i^*(t) \end{aligned}$$

Proof. See Bretti et al. (2006). □

According to Lemma 2.3, the separating shock emerges from the downstream boundary b_i of the link, and propagates towards the interior of the link. The speed of this separating shock is given by the Rankine-Hugoniot condition Evans (2010).

It is clear that as long as the separating shock remains in the interior of the link I_i, the upstream and downstream boundary conditions do not interact. Thus the exit of a link remains in the congested phase; while the entrance remains in the free flow phase. Consequently, the Riemann Solver in Theorem 2.1 is expressed entirely with exogenous parameters C_i and $\alpha_{1,i}$.

On the other hand, if the separating shock reaches either boundary, it becomes a latent shock. Two cases may arise.

i) The the shock reaches the exit, i.e. $x_i^*(t) = b_i$. In this case, the current link is dominated by free flow phase. The boundary condition at $x = a_i$ directly influences the boundary condition at $x = b_i$, in a way expressed by

$$q_i(t, b_i-) = q_i\left(t - \frac{L_i}{k_i}, a_i\right) \tag{2.10}$$

where L_i is the length of the link, k_i is the speed of forward wave propagation. See Figure 1 for an illustration.

ii) The shock reaches the entrance, i.e. $x_i^*(t) = a_i$. In this case, the current link is dominated by congested phase. The boundary condition at $x = b_i$ directly affects the boundary condition at $x = a_i$

$$q_i(t, a_i+) = q_i\left(t - \frac{L_i}{w_i}, b_i\right) \tag{2.11}$$

where w_i is the speed of backward wave propagation.

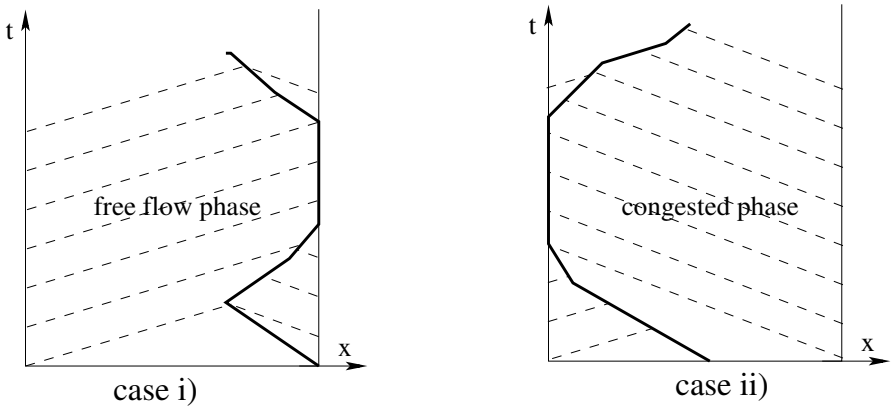

Figure 1: Example of latent separating shocks. Case i): the separating shock reaches the right (downstream) boundary. Case ii): the separating shock reaches the left (upstream) boundary.

In either case, the Riemann solver involves boundary flows (2.10), (2.11), which are endogenous. The next key step towards the link-based flow model is the detection of these two extreme cases. This can be done with a variational method called the Lax-Hopf formula (Aubin et al., 2008; Daganzo, 2005, 2006; Lax, 1957, 1973; Newell, 1993).

2.4 The variational approach for detecting latent shock

This section provides sufficient and necessary condition for the occurrence of the latent shock. The derivation is omitted for brevity, we refer the reader to Han et al. (2012) for a detailed discussion. Define for each link I_i the cumulative entering and exiting vehicle numbers

$$N_{i,up}(t) \doteq \int_0^t q_i(s, a_i)\, ds, \qquad N_{i,down} \doteq \int_0^t q_i(s, b_i)\, ds$$

Recall that the separating shock $x_i^*(\cdot) : [0, +\infty) \to [a_i, b_i]$ is a continuous curve in the $t-x$ domain. The following theorem provides sufficient and necessary conditions for the occurrence of the latent shock.

Theorem 2.4. *Let $N_i(\cdot,\cdot) : [0, +\infty) \times [a_i, b_i]$ be the unique viscosity solution to Hamilton-Jacobi equation (1.3) satisfying zero initial condition, upstream boundary condition $N_{i,up}(\cdot)$ and downstream boundary condition $N_{i,down}(\cdot)$. Then for all $t \geq 0$,*

$$x_i^*(t) \;=\; a_i \iff N_{i,up}(t) \;\geq\; N_{i,down}\left(t - \frac{L_i}{w_i}\right) + \rho_i^{jam} L_i \tag{2.12}$$

$$x_i^*(t) \;=\; b_i \iff N_{i,up}\left(t - \frac{L_i}{k_i}\right) \;\leq\; N_{i,down}(t) \tag{2.13}$$

Remark 2.5. *The significance of criteria (2.12)-(2.13) is that the two extreme cases can be detected without any computation within the link. This is because $N_{up}(\cdot)$, $N_{down}(\cdot)$ are determined completely by the boundary flows. Theorem 2.4 is the key ingredient of the LKWM, which allows the network model to be solved at the link level.*

Analytical properties of the LKWM pertaining to solution existence, uniqueness and well-posedness are provided in Han et al. (2012).

3 Traffic signal control problem based on the LKWM

In this section, the signal control problem is formulated with LKWM. We start with a single junction, with two incoming links I_1, I_2, and two outgoing links I_3, I_4 (Figure 2). Each link

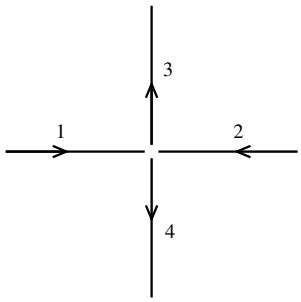

Figure 2: A signalized junction with two incoming links and two outgoing links.

is represented by a spatial interval $[a_i, b_i]$, $i = 1, 2, 3, 4$. The fundamental diagrams for each link is given by

$$f_i(\rho) \;=\; \begin{cases} k_i\,\rho & \rho \in [0,\, \rho_i^*] \\ -w_i\,(\rho - \rho_i^{jam}) & \rho \in (\rho_i^*,\, \rho_i^{jam}] \end{cases}, \qquad i \;=\; 1,\, 2,\, 3,\, 4$$

with $C_i \doteq k_i\,\rho_i^*$ being the flow capacity. We make the following assumptions
A1 The network is initially empty.
A2 Drivers arriving at the junction distribute on the outgoing roads according to some known coefficients:

$$A \;=\; \begin{pmatrix} \alpha_{1,3} & \alpha_{1,4} \\ \alpha_{2,3} & \alpha_{2,4} \end{pmatrix}$$

where α_{ij} denotes the percentage of traffic coming from link I_i that distributes to outgoing link I_j.

In the problem setting, the flows $\bar{q}_i(\cdot)$, $i = 1, 2$ entering links I_1 and I_2 are known. In practice, $\bar{q}_i(\cdot)$ can be measured at the entrance using fixed sensors such as loop-detectors.

3.1 Continuous-time formulation

In this section, we formulate the constraints of the system in continuous time. In Section 3.2 we will reformulate the system dynamic as linear constraints in discrete time using binary variables. Notice that the discussion in this section can be the building block for extension to networks with multiple intersections.

Let us fix the planning horizon $[0,\, T]$ for some fixed $T > 0$. Introducing the piecewise-constant control variables $u_i(\cdot) : [0,\, T] \to \{0,\, 1\}$, $i = 1,\, 2$, with the agreement that $u_i(t) = 0$ if the light is red for link I_i, and $u_i(t) = 1$ if the light is green for link I_i. It is convenient to use the following set of notations. For $i = 1,\, 2,\, 3,\, 4$.

$\bar{q}_i(\cdot)$,	the flow of cars entering link I_i,
$\hat{q}_i(\cdot)$,	the flow of cars exiting link I_i,
$\bar{r}_i(\cdot)$,	the binary variable that indicates the regime at $x = a_i+$,
$\hat{r}_i(\cdot)$,	the binary variable that indicates the regime at $x = b_i-$,
$\bar{q}_i^{max}(\cdot)$,	the maximum flow allowed to enter the link I_i,
$\hat{q}_i^{max}(\cdot)$,	the maximum flow allowed to exit the link I_i,
$N_{up,i}(\cdot)$,	the cumulative number of cars that have entered link I_i,
$N_{down,i}(\cdot)$,	the cumulative number of cars that have exited link I_i,
$u_i(\cdot)$,	the signal control variable for link I_i,

Theorem 3.1. *The dynamics at the junction (Figure 2) with signal control can be described by the following system of differential algebraic equations (DAE) with binary variables.*

$$\frac{d}{dt}N_{up,i}(t) = \bar{q}_i(t), \qquad \frac{d}{dt}N_{down,i}(t) = \hat{q}_i(t), \qquad i = 1,\, 2,\, 3,\, 4 \tag{3.14}$$

$$\bar{r}_i(t) = \begin{cases} 1, & \text{if} \quad N_{up,i}(t) \geq N_{down,i}\left(t - \frac{L_i}{w_i}\right) + \rho_i^{jam} L_i \\ 0, & \text{otherwise} \end{cases}, \qquad i = 1,\, 2,\, 3,\, 4 \tag{3.15}$$

$$\hat{r}_i(t) = \begin{cases} 0, & \text{if} \quad N_{up,i}\left(t - \frac{L_i}{k_i}\right) \leq N_{down,i}(t) \\ 1, & \text{otherwise} \end{cases}, \qquad i = 1,\, 2,\, 3,\, 4 \tag{3.16}$$

$$\bar{q}_i^{max}(t) = C_i + \bar{r}_i(t)\left(\hat{q}_i\left(t - \frac{L_i}{w_i}\right) - C_i\right), \qquad i = 1,\, 2 \tag{3.17}$$

$$\hat{q}_i^{max}(t) = \bar{q}_i\left(t - \frac{L_i}{k_i}\right) + \hat{r}_i(t)\left(C_i - \bar{q}_i\left(t - \frac{L_i}{k_i}\right)\right), \qquad i = 3,\, 4 \tag{3.18}$$

$$\hat{q}_i(t) = \begin{cases} 0, & \text{if} \quad u_i(t) = 0 \\ \min\left\{\hat{q}_i^{max}(t),\, \frac{\bar{q}_3^{max}(t)}{\alpha_{i,3}},\, \frac{\bar{q}_4^{max}(t)}{\alpha_{i,4}}\right\} & \text{if} \quad u_i(t) = 1 \end{cases}, \qquad i = 1,\, 2 \tag{3.19}$$

$$\bar{q}_k(t) = \alpha_{1,k}\,\hat{q}_1(t) + \alpha_{2,k}\,\hat{q}_2(t), \qquad k = 3,\, 4, \tag{3.20}$$

$$u_1(t) + u_2(t) = 1 \qquad \text{for all} \quad t \in [0,\, T] \tag{3.21}$$

Proof. (3.14) is by definition. For $i = 1, 2, 3, 4$, if the separating shock on link I_i reaches the entrance a_i (exit b_i), then the regime variable $\bar{r}_i = 1$ ($\bar{r}_i = 0$). Then (3.15)-(3.16) follows from Theorem 2.4.

The demand function $\hat{q}_i^{max}(\cdot)$ for incoming links and the supply function $\bar{q}_i^{max}(\cdot)$ for outgoing links are given by (2.9): for $i = 1,\, 2$, if $\hat{r}_i(t) = 1$, then $\hat{q}_i^{max}(t) = C_i$; otherwise if $\hat{r}_i(t) = 0$,

then according to (2.10),

$$\hat{q}_i^{max}(t) = q_i(t,\, b_i-) = q_i\left(t - \frac{L_i}{k_i},\, a_i\right) = \bar{q}_i\left(t - \frac{L_i}{k_i}\right)$$

This shows (3.18). One can similarly show (3.17) using (2.9) and (2.11).

For $i = 1,\, 2$, if $u_i(t) = 0$ which means the light is red, then the flow allowed through is zero, otherwise, it is given by(2.5). This proves (3.19).

(3.20) follows from the definition of the splitting parameters $\alpha_{i,k}$, $i = 1,\, 2$, $k = 3,\, 4$. (3.20) guarantees that at each time, there are one and only one incoming road that has green light. □

3.2 Discrete-time formulation

In this section, we present the discrete-time version of the optimization problem in Theorem 3.1. Let us introduce a few more notations for the convenience of our presentation. Consider a uniform time grid

$$0 \;=\; t^0 \;<\; t^1 \;\ldots\; <\; t^N \;=\; T, \qquad t^j - t^{j-1} \;=\; \delta t, \quad j \;=\; 1, \ldots, N$$

Throughout the rest of this article, we use superscript 'j' to denote the discrete value evaluated at time step t^j. In addition, we let $L_i/k_i = \Delta_i^f \delta t$, $L_i/w_i = \Delta_i^b \delta t$, $\Delta_i^f \in \mathbb{N}$, $\Delta_i^b \in \mathbb{N}$, $i = 1, 2, 3, 4$.

Approximating the numerical integration with rectangular quadratures, we write equality (3.15) and (3.16) in discrete time as

$$\begin{cases} \delta t \sum_{j=0}^{k-\Delta_i^b} \hat{q}_i^j - \delta t \sum_{j=0}^{k} \bar{q}_i^j + \rho^{jam} L_i \;\le\; \mathcal{M}\,(1 - \bar{r}_i^k) \\ \delta t \sum_{j=0}^{k-\Delta_i^b} \hat{q}_i^j - \delta t \sum_{j=0}^{k} \bar{q}_i^j + \rho^{jam} L_i \;\ge\; -\mathcal{M}\,\bar{r}_i^k + \varepsilon \end{cases} \qquad \Delta_i^b \;\le\; k \;\le\; N, \quad i \;=\; 1,\, 2,\, 3,\, 4 \tag{3.22}$$

$$\begin{cases} \delta t \sum_{j=0}^{k-\Delta_i^f} \bar{q}_i^j - \delta t \sum_{j=0}^{k} \hat{q}_i^j \;\le\; \mathcal{M}\,\hat{r}_i^k \\ \delta t \sum_{j=0}^{k-\Delta_i^f} \bar{q}_i^j - \delta t \sum_{j=0}^{k} \hat{q}_i^j \;\ge\; \mathcal{M}\,(\hat{r}_i^k - 1) + \varepsilon \end{cases} \qquad \Delta_i^f \;\le\; k \;\le\; N, \quad i \;=\; 1,\, 2,\, 3,\, 4 \tag{3.23}$$

where $\bar{r}_i^k,\, \hat{r}_i^k \in \{0,\, 1\}$. $\mathcal{M} \in \mathbb{R}_+$ is a sufficiently large number, $\varepsilon \in \mathbb{R}_+$ is a sufficiently small number. Constraints (3.22) and (3.23) determines the regime variables associated with the two boundaries of each link. Once the flow phases are determined, the demand and supply functions (3.17), (3.18) are re-written in discrete time as

$$\begin{cases} C_i - \mathcal{M}\,\bar{r}_i^j \;\le\; \bar{q}_i^{max,j} \;\le\; C_i \\ \hat{q}_i^{j-\Delta_i^b} - \mathcal{M}\,(1 - \bar{r}_i^j) \;\le\; \bar{q}_i^{max,j} \;\le\; \hat{q}_i^{j-\Delta_i^b} + \mathcal{M}\,(1 - \bar{r}_i^j) \end{cases} \qquad i \;=\; 1,\, 2,\, 3,\, 4 \tag{3.24}$$

$$\begin{cases} C_i + \mathcal{M}\,(\hat{r}_i^j - 1) \;\le\; \hat{q}_i^{max,j} \;\le\; C_i \\ \bar{q}_i^{j-\Delta_i^f} - \mathcal{M}\,\hat{r}_i^j \;\le\; \hat{q}_i^{max,j} \;\le\; \bar{q}_i^{j-\Delta_i^f} + \mathcal{M}\,\hat{r}_i^j \end{cases} \qquad i \;=\; 1,\, 2,\, 3,\, 4 \tag{3.25}$$

Next, let us re-formulate (3.19). Introducing dummy variables ζ_1^j, ζ_2^j, $1 \le j \le N$, such that

$$\zeta_i^j \;=\; \min\left\{\hat{q}_i^{max,j},\, \frac{\bar{q}_3^{max,j}}{\alpha_{i,3}}\, \frac{\bar{q}_4^{max,j}}{\alpha_{i,4}}\right\} \qquad i \;=\; 1,\, 2 \tag{3.26}$$

Then the discrete-time version of (3.19) can be readily written as

$$\begin{cases} 0 \;\le\; \hat{q}_i^j \;\le\; \mathcal{M}\, u_i^j \\ \zeta_1^j + \mathcal{M}\,(u_i^j - 1) \;\le\; \hat{q}_i^j \;\le\; \zeta_1^j \end{cases} \qquad i \;=\; 1,\, 2, \quad j \;=\; 1, \ldots, N \tag{3.27}$$

In order to write (3.26) as linear constraints, one could write it as three "less or equal" statements, which is simple but bear the potential limitation of traffic holding. Instead, one may introduce additional binary variables ξ_i^j, η_i^j and real variables β_i^j for $i = 1,\, 2$, $j = 1, \ldots, N$, such that (3.26) can be accurately formulated as

$$\begin{cases} \bar{q}_3^{max,j}/\alpha_{i,3} - \mathcal{M}\,\xi_i^j \;\le\; \beta_i^j \;\le\; \bar{q}_3^{max,j}/\alpha_{i,3} \\ \bar{q}_4^{max,j}/\alpha_{i,4} - \mathcal{M}\,(1-\xi_i^j) \;\le\; \beta_i^j \;\le\; \bar{q}_4^{max,j}/\alpha_{i,4} \\ \hat{q}_i^{max,j} - \mathcal{M}\,\eta_i^j \;\le\; \zeta_i^j \;\le\; \hat{q}_i^{max,j} \\ \beta_i^j - \mathcal{M}\,(1-\eta_i^j) \;\le\; \zeta_i^j \;\le\; \beta_i^j \end{cases} \qquad i \;=\; 1,\, 2 \tag{3.28}$$

Finally, we have the obvious relations

$$\bar{q}_k^j(t) \;=\; \alpha_{1,k}\hat{q}_1^j(t) + \alpha_{2,k}\hat{q}_2^j(t) \qquad k \;=\; 3,\, 4, \qquad j \;=\; 1, \ldots, N \tag{3.29}$$

and

$$u_1^j + u_2^j \;=\; 1 \qquad j \;=\; 1, \ldots, N \tag{3.30}$$

The proposed MILP formulation of signal control problem is summarized by (3.22)-(3.25) and (3.27)-(3.30). This formulation captures many desirable features of vehicular flow on networks such as physical queues, spill back, vehicle turning, and shock formation and propagation (although not explicitly). The signal control allows time-varying cycle length and splits, as well as the utilization of real-time information of traffic flows.

3.3 Bound the separating shock

At the end of this section, we discuss an additional linear constraint that ensures that the congested phase on each link is bounded. Such condition is closely related to travel delay: if the congested phase remain bounded, there will be a reasonable upper bound for the travel time of each driver. Let us consider a single link $[a,\, b]$. If we wish to bound the separating shock within the interval $[c,\, b]$ for some $a < c < b$, this means (with the same notation as before) that

$$\int_0^t q(s,\, c)\, ds \;\le\; \int_0^{t-\frac{b-c}{w}} \hat{q}(s)\, ds + \rho^{jam}\,(b-c) \tag{3.31}$$

(3.31) follows by applying (2.12) to the interval $[c,\, b]$. It is helpful to notice that since $[a,\, c]$ remains in the free flow phase, $q(s,\, c)$ must be equal to $\bar{q}\left(s - \frac{c-a}{k}\right)$. Thus the condition for the congested phase to remain in $[c,\, b]$ becomes

$$N_{up}\left(t - \frac{c-a}{k}\right) \;\le\; N_{down}\left(t - \frac{b-c}{w}\right) + \rho^{jam}\,(b-c) \tag{3.32}$$

Condition (3.32) can be easily written as linear constraint in discrete time. If one wishes to include (3.32) in the objective function instead of using it as a constraint, he/she may simply minimize the difference between the left and right hand sides of (3.32).

4 Emission-related side constraints

In this section, we propose an emission constraint that employs robust optimization to handle parameter errors. It is intuitively clear that the total emission amount on an link of interest is highly correlated with the total number of vehicles on the same link. Therefore, it is natural to approximate a relationship between the number of vehicles on link I_i, denoted by $N_i(t) = N_{i,up}(t) - N_{i,down}(t)$, and the aggregated emission rate ϵ_i on link I_i. We assume such a relationship is expressible by a polynomial with power $\mathcal{L}$.

$$\epsilon_i(\boldsymbol{a},\ N_i(t)) = \sum_{l=0}^{\mathcal{L}} a_l (N_i(t))^l = \sum_{l=0}^{\mathcal{L}} a_l (N_{i,up}(t) - N_{i,down}(t))^l = \boldsymbol{a}^T \boldsymbol{N}_i(t) \tag{4.33}$$

where $\boldsymbol{a} = (a_0, a_1, \ldots, a_{\mathcal{L}})^T$, $\boldsymbol{N}_i(t) = \big((N_i(t))^0, \ldots, (N_i(t))^{\mathcal{L}}\big)$. With the these notations, the emission side constraint is expressed as

$$\int_0^T \epsilon_i(\boldsymbol{a},\ N_i(t))dt \ \leq\ E_i \tag{4.34}$$

Such a constraint requires that the emission amount is below some critical value E_i for any link I_i. Notice that such constraints can be easily transformed to address other types of environmental considerations including

- the total emission amount on the entire network is bounded by some given value
- the differences among total emissions of all links are bounded or minimized

The last one ensures that no link (or the neighborhood near that link) suffers much more than other links in terms of air quality. Such an issue is identified by Benedek and Rilett (1998) as *environmental equity*.

4.1 Side constraints considering errors

In (4.34), approximation is involved in the relationship between the vehicle number and the emission amount. In order to ensure the emission constraint is satisfied even with parameter errors, the following robust constraint is of our interest, instead:

$$\int_0^T \hat{\epsilon}_i(\boldsymbol{a}(t),\ N_i(t))dt \ \leq\ E_i \qquad \text{for all } \boldsymbol{a}(t) \in \eta_{a,i} \tag{4.35}$$

where by allowing the coefficient $\boldsymbol{a}(t) = (a_l(t) : 0 \leq l \leq \mathcal{L})$ to be varied over time, we have

$$\hat{\epsilon}_i(\boldsymbol{a}(t),\ N_i(t)) = \sum_{k=0}^{\mathcal{L}} a_l(t)(N_i(t))^l \tag{4.36}$$

and $\eta_{a,i}$ is specified as a budget-like uncertainty similar to Atamtürk and Zhang (2007) given as:

$$\eta_{a,i} \ =\ \left\{ a :\ L_l \leq a_l(t) \leq U_l,\ \text{for all } 0 \leq l \leq \mathcal{L},\ \sum_{l=1}^{\mathcal{L}} \int_0^T a_l(t)\, dt \ \leq\ \frac{T \sum_{l=1}^{\mathcal{L}} U_l}{\theta} \right\} \tag{4.37}$$

where $\boldsymbol{L} := (L_l : 0 \leq l \leq \mathcal{L}) \in \mathbb{R}^{\mathcal{L}+1}$ and $\boldsymbol{U} := (U_l : 0 \leq l \leq \mathcal{L}) \in \mathbb{R}^{\mathcal{L}+1}$ are lower and upper bounds. $\theta \in \mathbb{R}_+$ is a scalar that adjusts the conservatism. Such a constraint corresponds to

the notion of robust optimization in the sense that the emission amount on link i is restricted to an upper bound with any possible realization of the parameter $\boldsymbol{a}(t)$.

Increasing the value of θ leads to a less conservative model. In the most conservative case, $0 \leq \theta \leq 1$, the last constraint in (4.51) is out of effect.

4.2 Discretization and explicit reformulation

Constraint (4.35) can be time-discretized into the following form

$$\sum_{k=0}^{N}\sum_{l=0}^{\mathcal{L}} a_{l,k}(N_{i,k})^l \delta t \;\leq\; E_i \qquad \text{for all } \hat{\boldsymbol{a}} \in \hat{\eta}_{a,i} \tag{4.38}$$

where $\hat{\boldsymbol{a}} := (a_{l,k} :\; 0 \leq l \leq \mathcal{L},\; 0 \leq k \leq N)$ with the uncertainty set discretized as:

$$\hat{\eta}_{a,i} = \left\{ \hat{\boldsymbol{a}} : L_l \leq a_{l,k} \leq U_l,\; \text{for all } 0 \leq l \leq \mathcal{L},\; 0 \leq k \leq N,\; \sum_{k=0}^{N}\sum_{l=1}^{\mathcal{L}} a_{l,k}\delta t \leq \frac{T\sum_{l=1}^{\mathcal{L}} U_l}{\theta} \right\} \tag{4.39}$$

Such a constraint is in fact a semi-infinite constraint with an infinite index set, which is not directly computable. The following theorem gives its computable reformulation.

Theorem 4.1. *Assume that the robust program has a nonempty feasible region. If $\hat{\eta}_{a,i}$ is nonempty, the semi-infinite constraint (4.38) is equivalent to the following set of constraints:*

$$\sum_{l=1}^{\mathcal{L}}\sum_{k=0}^{N} d_{1,l,k}U_l - \sum_{l=1}^{\mathcal{L}}\sum_{k=0}^{N} d_{2,l,k}L_l + \frac{T\sum_{l=1}^{\mathcal{L}} U_l}{\theta}\, d_3 + \sum_{k=0}^{N} U_0\, \delta t \;\leq\; E_i \tag{4.40}$$

$$s.t.\;\; d_{1,l,k} - d_{2,l,k} + d_3 = (N_{i,k})^l \delta t \qquad \textit{for all}\;\; 1 \leq l \leq \mathcal{L}, \qquad 0 \leq k \leq N \tag{4.41}$$

$$d_{1,l,k},\; d_{2,l,k},\; d_3 \geq 0 \qquad \textit{for all}\;\; 1 \leq l \leq \mathcal{L}, \qquad 0 \leq k \leq N \tag{4.42}$$

Proof. Constraint (4.38) can be trivially rewritten as:

$$\max_{\hat{\boldsymbol{a}} \in \hat{\eta}_{a,i}} \sum_{k=0}^{N}\sum_{l=1}^{\mathcal{L}} a_{l,k}(N_{i,k})^l \delta t + \max_{\hat{\boldsymbol{a}} \in \hat{\eta}_{a,i}} \sum_{k=0}^{N} a_{0,k}(N_{i,k})^0 \delta t \leq E_i, \tag{4.43}$$

which is equivalent to

$$\max_{\hat{\boldsymbol{a}} \in \hat{\eta}_{a,i}} \sum_{k=0}^{N}\sum_{l=1}^{\mathcal{L}} a_{l,k}(N_{i,k})^l \delta t \leq E_i - \sum_{k=0}^{N} U_0 \delta t. \tag{4.44}$$

The evaluation of the constraint function involves solving a parametric problem of the form:

$$\max_{\hat{\boldsymbol{a}}} \sum_{k=0}^{N}\sum_{l=1}^{\mathcal{L}} a_{l,k}(N_{i,k})^l \delta t \tag{4.45}$$

$$s.t.\;\; \hat{\boldsymbol{a}} \in \hat{\eta}_{a,i} \tag{4.46}$$

where we treat the $N_{i,k}$ as a constant parameter. Program (4.45)-(4.46) has a dual problem of the following:

$$\min_{\boldsymbol{D}} \sum_{l=1}^{\mathcal{L}} \sum_{k=0}^{N} d_{1,l,k} U_l - \sum_{l=1}^{\mathcal{L}} \sum_{k=0}^{N} d_{2,l,k} L_l + \frac{T \sum_{l=1}^{\mathcal{L}} U_l}{\theta} d_3 \tag{4.47}$$

$$s.t. \;\; d_{1,l,k} - d_{2,l,k} + d_3 = (N_{i,k})^l \delta t \qquad \text{for all } \; 1 \le l \le \mathcal{L}, \qquad 0 \le k \le N \tag{4.48}$$

$$d_{1,l,k},\; d_{2,l,k},\; d_3 \ge 0 \qquad \text{for all } \; 1 \le l \le \mathcal{L}, \qquad 0 \le k \le N \tag{4.49}$$

Under the assumption of nonempty feasible region and nonempty $\hat{\eta}_{a,i}$, by noticing the compactness of $\hat{\eta}_{a,i}$, the primal program (4.45)-(4.46) and the dual program (4.47)-(4.49) have finite solutions and zero duality gap. Moreover, by duality, the objective value of any feasible solution to the dual problem (4.47)-(4.49) provides upper bounds to the primal problem (4.45)-(4.46). Therefore, if there exist $d_{1,l,k}$, $d_{2,l,k}$ and d_3 such that (4.48)-(4.49) are satisfied, then, if $N_{i,k}$ satisfy (4.40), then (4.44) is satisfied.

On the other hand, if there exists $N_{i,k}$ that satisfy (4.44), then the objective value of the optimal solution to the parametric problem (4.45)-(4.46) is bounded above and thus there exists $d_{1,l,k}$, $d_{2,l,k}$ and d_3 such that (4.48)-(4.49) are satisfied. Thus, the equivalence of interest is proved. □

Remark 4.2. *The above theorem is based on the discussions in Ben-Tal and Nemirovski (1999) and Bertsimas et al. (2011a). With such reformulation the original semi-infinite constraint is now computable with standard nonlinear programming techniques. We refer to the reformulated program with constraints (4.40)-(4.42) as the robust counterpart to the original robust problem.*

4.3 A special case when the relationship is linear

In an desirably simple case where a linear correlation between the vehicle number and the emission amount is detected, the robust problem can be considerably simplified. In such a case, we can reduce (4.35) to

$$\begin{aligned} &\int_0^T [a_1(t)\,(N_i(t)) + a_0(t)]dt \\ &\quad = \int_0^T [a_1(t)\,(N_{up,i}(t) - N_{down,i}(t)) + a_0(t)]dt \le E_i, \quad \text{for all } (a_0^T(t), a_1^T(t))^T \in \eta_{a,i} \end{aligned} \tag{4.50}$$

where

$$\eta_{a,i} = \left\{ (a_1^T(t),\, a_2^T(t))^T :\; L_l \le a_l(t) \le U_l,\; l = 0,\, 1,\; \int_0^T a_1(t)dt \le \frac{TU_1}{\theta} = \frac{(N+1)\delta t U_1}{\theta} \right\} \tag{4.51}$$

With discretization, we have the following formulation,

$$\sum_{k=0}^{N} \left[a_{1,k} \left(\delta t \sum_{j=0}^{k} \bar{q}_i^j - \delta t \sum_{j=0}^{k} \hat{q}_i^j \right) + a_{0,k} \right] \delta t \;\le\; E_i \qquad \text{for all } (a_0,\; a_1) \in \hat{\eta}_{a,i} \tag{4.52}$$

where $a_0 := (a_{0,k} : 0 \le k \le N)$, $a_1 := (a_{1,k} : 0 \le k \le N)$. The uncertainty set is given as

$$\hat{\eta}_{a,i} = \left\{ (a_0,\, a_1) : L_0 \le a_{0,k} \le U_0,\ L_1 \le a_{1,k} \le U_1,\ \text{for all } 0 \le k \le N,\ \sum_{k=0}^{N} a_{1,k} \le \frac{(N+1)U_1}{\theta} \right\} \tag{4.53}$$

With the above uncertainty set, we have the following reformulation:

$$\sum_{k=0}^{N} U_1 d_{1,k} - \sum_{k=0}^{N} L_1 d_{2,k} + \frac{(N+1)U_1}{\theta} d_3 + (N+1)U_2\, \delta t \ \le\ E_i \tag{4.54}$$

$$d_3 + d_{1,k} - d_{2,k} \ =\ \delta t^2 \sum_{j=0}^{k} \bar{q}_i^j - \delta t^2 \sum_{j=0}^{k} \hat{q}_i^j \qquad \text{for all } k = 0, 1, ..., N \tag{4.55}$$

$$d_{1,k},\ d_{2,k},\ d_3 \ \ge\ 0 \qquad \text{for all } k = 0, 1, ..., N \tag{4.56}$$

where $d_{1,k}$, $d_{2,k}$, d_3 are dual variables. In such a linear case, the semi-infinite constraint (4.50) is reformulated into a set of linear constraints. Thus, the robust problem can be solved efficiently with the state-of-the-art MIP solvers.

5 An empirical relationship between link occupancy and aggregated emission rate

In order to apply the robust optimization techniques from Section 4, we propose to investigate the correlation between the link occupancy and the *aggregated emission rate* (AER). Recall that for link i, $N_{up,i}(t)$, $N_{down,i}(t)$ denotes the cumulative entering and exiting curves, respectively. We let $\rho_i(t,\, x) : [0,\, T] \times [a_i,\, b_i] \to [0,\, \rho_i^{jam}]$ be the unique entropy solution to the conservation law (1.1) which is consistent with the boundary conditions given by $N_{i,up}(t)$, $N_{i,down}(t)$. In addition, let us assume a speed-based emission model of the form

$$e_t \ =\ \Upsilon(v) \tag{5.57}$$

where e_t (in gram / hour) is the amount of emission per unit time, v (in mile / hour) is the vehicle velocity. The aggregated emission rate on the link i at time instance t is the following.

$$\mathrm{AER}(t) \ \doteq\ \int_{a_i}^{b_i} \rho(t,\, x) \Upsilon\big(v(t,\, x)\big)\, dx \qquad t \in [0,\, T] \tag{5.58}$$

where $v(t,\, x)$ is expressed in terms of $\rho(t,\, x)$ via the fundamental diagram.

We are interested to know if there is any correlation between $AER(t)$, the aggregated emission rate on a link level, and $N(t)$, the current link occupancy. If such a relationship exists, at least empirically, then we can apply the robust optimization techniques from Section 4 to obtain a tractable mathematical program.

To this end, we employ the speed-based emission model TRANSYT-7F, which was used by Benedek and Rilett (1998); Penic and Upchurch (1992); Nagurney et al. (2010)

$$e_x \ =\ 26.3009 \cdot \frac{\exp(0.009928v)}{v} \tag{5.59}$$

where v (in mile / hour) is the vehicle speed, e_x (in gram per mile) is the amount of emission per unit travel distance. It is useful to notice that the amount of emission per unit time, denoted by e_t (in gram / hour), is determined by

$$e_t \ =\ e_x \cdot v \ =\ 26.3009 \cdot \exp(0.009928v) \tag{5.60}$$

5.1 Numerical simulation

We conduct a sequence of simulations to obtain sufficient data for the study of correlations between the quantities of interest. In order to proceed, we consider a single link with the following parameters.

$$k = 30 \text{ (mile / hour)}, \quad w = 10 \text{ (mile / hour)}, \quad \rho^{jam} = 400 \text{ (vehicle / mile)}, \quad L = 0.3 \text{ (mile)}.$$

This choice of parameters implies a flow capacity of $C = 3000$ (vehicle / hour). We use a time grid

$$0 \;=\; t^0 \;<\; t^1 \;\ldots\; <\; t^N \;=\; T \qquad t^j - t^{j-1} \;=\; \delta t, \quad j \;=\; 1, \ldots, N$$

As usual, the superscript 'j' indicates the discrete value at the j^{th} grid point. We assume D^j (S^j), $j = 0, \ldots, N$ are the time-varying demand (supply) at the entrance (exit) of the link. In other words, D^j is the number of cars that need to enter the link during the j^{th} time interval; whereas S^j is the maximum number of cars that can advance to the downstream link in the j^{th} time interval. In addition, we let each D^j and S^j be uniformly distributed random variables whose ranges will be specified below.

We partition the link into cells,

$$a \;=\; x_0 \;<\; x_1 \;\leq\; \ldots \;\leq\; x_M \;=\; b \qquad x_l - x_{l-1} \;=\; \delta x, \quad l \;=\; 1, \ldots, M.$$

Then we use the Godunov scheme introduced by Godunov (1959) to compute the spatial distribution of density ρ_i^j, $i = 0, \ldots, M$ at each time interval $[t_j, t_{j+1}]$. The aggregated emission rate (5.58) is then approximated by

$$AER^j \;=\; \delta x \sum_{i=0}^{M} \rho_i^j \Upsilon\big(v_i^j\big), \qquad j \;=\; 0, \ldots, N$$

where v_i^j's are easily calculated from the fundamental diagram, given ρ_i^j's. Finally, the link occupancy at the j^{th} interval is given by $N^j = \delta x \sum_{i=0}^{M} \rho_i^j$.

We consider four different scenarios of demand D^j and supply S^j of the link. Recall $C = 3000$ (vehicle/hour) is the flow capacity of the link.

1. $\big\{D^j\big\}_{j=0}^N$ are uniformly distributed in $[0,\, C/2]$; $\big\{S^j\big\}_{j=0}^N$ are uniformly distributed in $[0,\, C]$.
2. $\big\{D^j\big\}_{j=0}^N$ are uniformly distributed in $[0,\, C/2]$; $\big\{S^j\big\}_{j=0}^N$ are uniformly distributed in $[0,\, C/2]$.
3. $\big\{D^j\big\}_{j=0}^N$ are uniformly distributed in $[0,\, C]$; $\big\{S^j\big\}_{j=0}^N$ are uniformly distributed in $[0,\, C]$.
4. $\big\{D^j\big\}_{j=0}^N$ are uniformly distributed in $[0,\, C]$; $\big\{S^j\big\}_{j=0}^N$ are uniformly distributed in $[0,\, C/2]$.

Each scenario contains 2000 samples. The scatter plots of aggregated emission rate AER verses link occupancy are shown in Figure 3 - 6, corresponding to scenarios 1 - 4, respectively. The four scenarios are plotted together in Figure 7.

The choices of demand and supply imply that scenarios 1- 4 are expected to have increased congestion levels. In the first scenario, the traffic is almost always in free flow phase, this means the travel speed is a constant, due to the triangular fundamental diagram. Thus the total emission rate is proportional to the number of cars (see Figure 3). In both the second

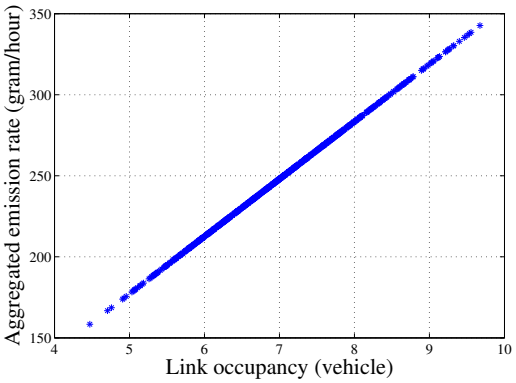

Figure 3: Scenario 1.

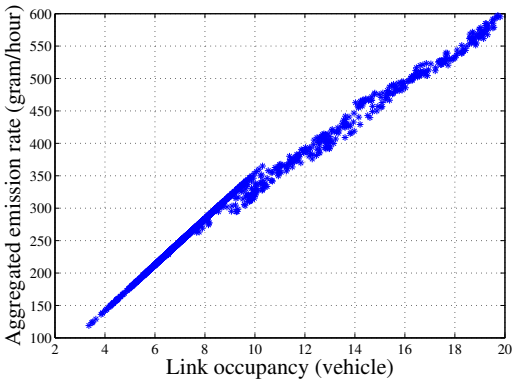

Figure 4: Scenario 2.

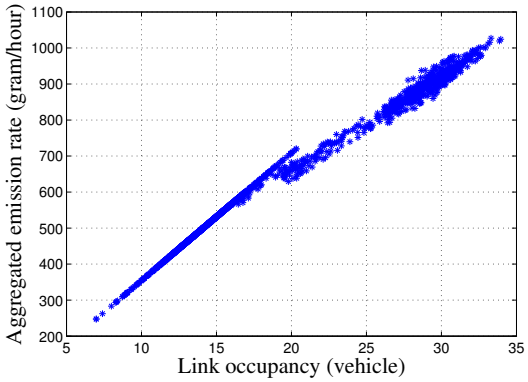

Figure 5: Scenario 3.

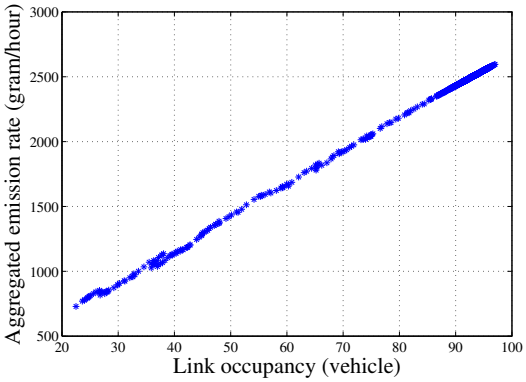

Figure 6: Scenario 4.

and third scenarios, the demand and the supply are comparable. Therefore we observe similar patterns in Figure 4 and 5, where the straight line corresponds to the case where the link is dominated by the free flow phase; and the more scattered portions imply the presence of both free flow phase and congested phase. In the last scenario, the link is dominated by the congested phase, yet the data still imply a linear relationship between link occupancy and aggregated emission rate.

With simple linear regression method, we may approximate the relationship between link occupancy and AER with an affine function. We can also obtain the lower and upper bounds for the coefficients, as shown in (4.53).

6 Numerical example

In this section, we consider the network consisting of two signalized junctions as shown in Figure 8. Assume that the inflow at the beginning of links I_1, I_2 and I_3 are given. These quantities can be measured, for example by loop detectors. We also fix the upper bound E_i for each link I_i so that the total emission cannot exceeds such an upper bound. The adaptive signal control problem with emission side constraints is then formulated as a mixed integer linear program with constraints (3.22)-(3.25), (3.27)-(3.30) and (4.54)-(4.56).

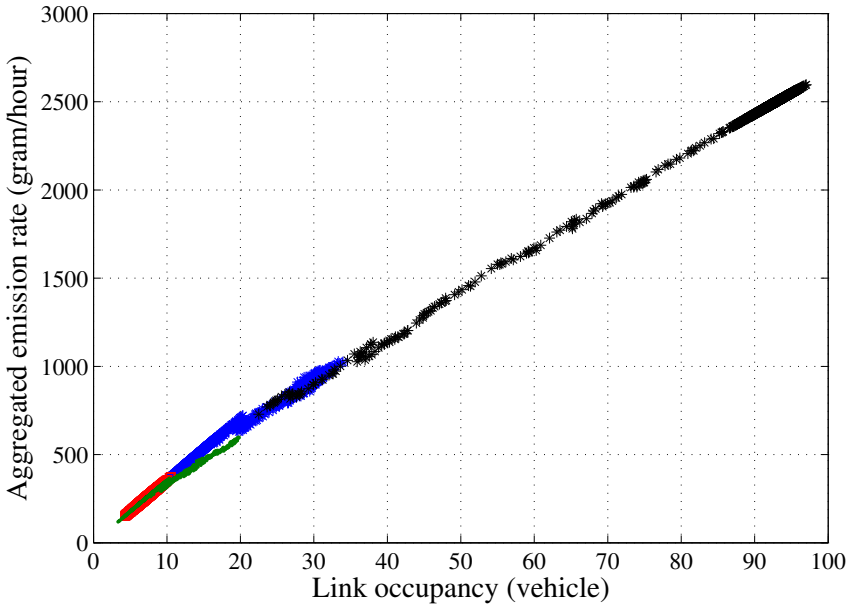

Figure 7: Scatter plot of link occupancy verses AER.

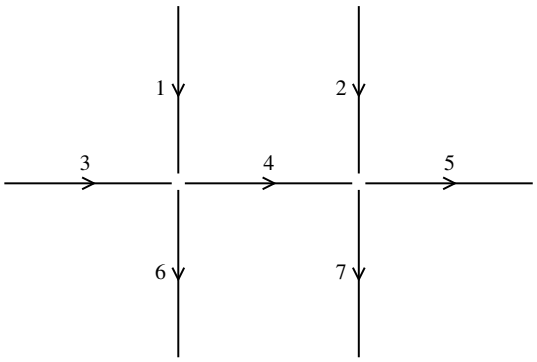

Figure 8: Test network with two signalized intersections.

6.1 Numerical setting

We assume the same fundamental diagram for all the links, that is, for $i = 1, \ldots, 7$,

$$k_i = 30 \text{ mile/hour}, \quad w_i = 10 \text{ mile/hour}, \quad \rho_i^{jam} = 400 \text{ vehicle/mile}, \quad C_i = 3000 \text{ vehicle/hour}$$

The lengths of all links are set equally to be 0.3 miles. We choose a time grid of 30 intervals and a time step of 0.005 hour (18 seconds). The flow entering links I_1, I_2, I_3 are chosen to be time-varying functions whose value at each time interval is randomly generated between 0 veh/hour and 3000 veh/hour. In addition, to ensure the performance of our signal control strategy, we include further constraint that the congested phase must never reach the upstream boundary of each link, i.e. no spillback occurs. This is done by invoking a special case of constraint (3.32), with $c = a$.

Regarding environmental consideration, we impose the constraints that the total emission emanating from each link, $I_1, \ldots, I_4$, does not exceed 150 grams.

The MBIP is solved with ILOG Cplex 12.1.0, which runs with Intel Xeon X5675 Six-Core 3.06 GHz processor provided by the Penn State Research Computing and Cyberinfrastructure.

6.2 Numerical results

In order to have a clear visualization of the optimal signal strategy and the separating shocks on each link, we use the boundary datum $\bar{q}_i$, $\hat{q}_i$, $i = 1, 2, 3, 4$ obtained from the optimal solution to construct solutions to the Hamilton-Jacobi equation (1.3), using the Lax-Hopf formula. For the H-J equation, the separating shock no longer represents discontinuity, rather, it is displayed as a 'kink' (discontinuity in the first derivative). The Moskowitz functions $N(t, x)$ for links I_3, I_4 are shown in Figure 10 and 12, respectively. The Moskowitz functions for link I_2, I_4 viewed from a different angle is shown in Figure 11 and 12.

One can clearly observe, in each figure, two types of characteristics: forward ones and backward ones. The mutual boundaries of the two regimes are identified as the location of the separating shock waves, which we managed to deal with implicitly using the variational method. It is also clear that the congested regions never reaches the left boundaries, indicating that vehicle spillbacks never occur.

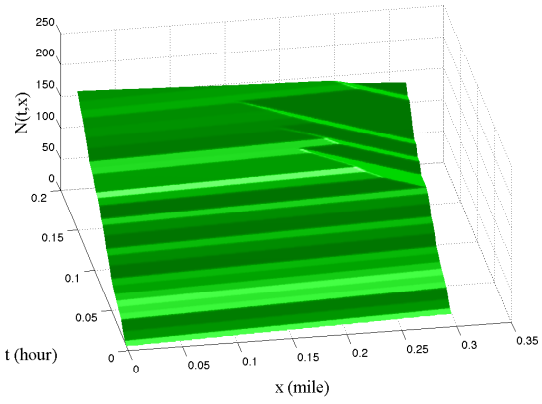

Figure 9: Moskowitz function of link I_1.

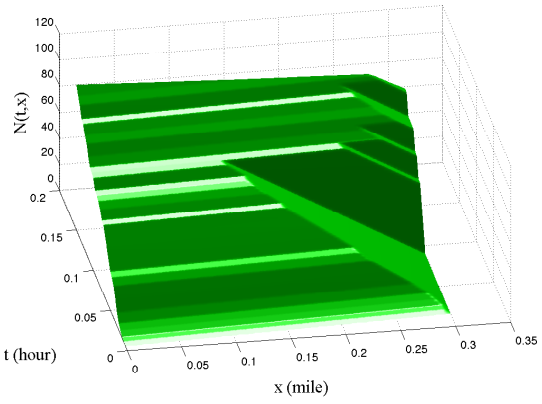

Figure 10: Moskowitz function of link I_3.

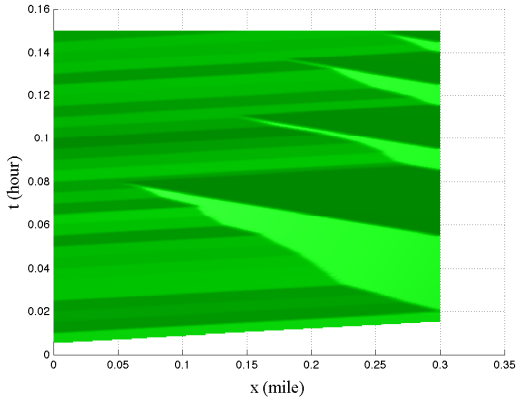

Figure 11: Moskowitz function of link I_2.

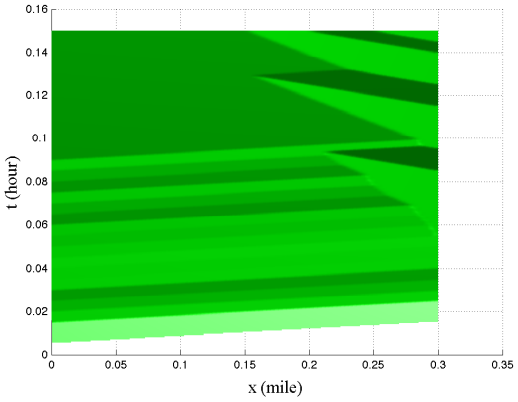

Figure 12: Moskowitz function of link I_4.

We show the emission results in Table 1 below. For comparison purpose, we also compute

the optimal signal strategy without any emission-related constraints. It can be seen from Table 1 that pollutants are mainly concentrated on links I_2 and I_3 without emission constraints; and the emission amount on I_2 exceeds 150 grams. On the other hand, by considering emissioin constraints and using RO approach, we managed to control the emission amount on all four links to below 150 grams, even under the worst case.

	Link 1	Link 2	Link 3	Link 4
Without emission constraints	99.9	161.4	144.7	99.5
With emission constraints and RO	109.8	149.5	134.8	111.4

Table 1: Link-specific total emissions (in gram). The values are computed according to the worse case over all uncertainty sets of the parameters.

7 Conclusion

This paper proposes a novel robust optimization approach to address emission-related side constraints, in the formulation of a link-based mathematical programming problem for traffic signal control. We take advantage of an empirical relationship between the aggregated emission rate and the link occupancy, and treat the error as modeling uncertainty which is then handled by the RO technique. Such a computational apparatus turns the otherwise non-convex emission side constraints into explicit and tractable forms.

Further research is underway to 1) validate the linear or polynomial relationships among certain macroscopic traffic quantities, from both statistical and analytical points of view; 2) reveal additional correlations between aggregated emission rate and certain traffic quantities that may facilitate the computational efficiency and tractability of the math programs; and 3) develop heuristic algorithms and computational paradigms for real-time and large-scale deployment.

References

Atamtürk, A., Zhang, M., 2007. Two-stage robust network flow and design under demand uncertainty. Operations Research 55 (4), 662-673.

Aubin, J.P., Bayen, A.M., Saint-Pierre, P., 2008. Dirichlet problems for some Hamilton-Jacobi equations with inequality constraints. SIAM Journal on Control and Optimization 47 (5), 2348-2380.

Bandi, C., D. Bertsimas. 2012. Tractable stochastic analysis in high dimensions via robust optimization. Mathematical Programming, Series B. DOI: 10.1007/s10107-012-0567-2

Ben-Tal, A., A. Nemirovski. 1998. Robust convex optimization. Mathematics of Operations Research 23(4) 769–805.

Ben-Tal, A., A. Nemirovski. 1999. Robust solutions of uncertain linear programs. Operations Research Letters 25(1) 1–14.

Ben-Tal, A., A. Nemirovski. 2000. Robust solutions of linear programming problems contaminated with uncertain data. Mathematical Programming 88 411-421.

Benedek, C.M., Rilett, L.R., 1998. Equitable traffic assignment with environmental cost function. Journal of Transportation Engineering, 124, 1622.

Bertsimas, D., D.B. Brown, C. Caramanis. 2011a. Theory and applications of robust optimization. SIAM Review 53(3) 464–501.

Bertsimas, D., Chang, A., Rudin, C., 2011. Integer optimization methods for supervised ranking. Available online at http://hdl.handle.net/1721.1/67362.

Bertsimas, D., Gamarnik, D., Rikun, A.A., 2011b. Performance analysis of queueing networks via robust optimization. Operations Research 59(2) 455–466.

Bertsimas, D., M. Sim. 2004. Price of robustness. Operations Research 52(1) 35–53.

Bressan, A., 2000. Hyperbolic Systems of Conservation Laws.The One Dimensional Cauchy Problem. Oxford University Press.

Bretti, G., Natalini, R., Piccoli, B., 2006. Numerical approximations of a traffic flow model on networks, Networks and Heterogeneous Media 1, 57-84.

Claudel, C.G., Bayen, A.M., 2010. Lax-Hopf Based Incorporation of Internal Boundary Conditions Into Hamilton-Jacobi Equation. Part I: Theory. IEEE Transactions on Automatic Control 55 (5), 1142-1157.

Coclite, G.M., Garavello, M., Piccoli, B., 2005. Traffic flow on a road network. SIAM Journal on Mathematical Analysis 36 (6), 1862-1886.

Daganzo, C.F., 1994. The cell transmission model. Part I: A simple dynamic representation of highway traffic. Transportation Research Part B 28 (4), 269-287.

Daganzo, C.F., 1995. The cell transmission model. Part II: Network traffic. Transportation Research Part B 29 (2), 79-93.

Daganzo, C.F., 2005. A variational formulation of kinematic waves: basic theory and complex boundary conditions. Transportation Research Part B 39 (2), 187-196.

Daganzo, C.F., 2006. On the variational theory of traffic flow: well-posedness, duality and application. Network and heterogeneous media 1 (4), 601-619.

Evans, L.C., 2010. Partial Differential Equations. Second edition. American Mathematical Society, Providence, RI.

Ferrari, P., 1995. Road pricing and network equilibrium. Transportation Research Part B 29 (5), 357-372.

Garavello, M., Piccoli, B., 2006. Traffic Flow on Networks. Conservation Laws Models. AIMS Series on Applied Mathematics, Springfield, Mo..

Gartner, N.H., 1983. OPAC: a demand-responsive strategy for traffic signal control. Transportation Research Record 906, 75-81.

Godunov, S.K., 1959. A difference scheme for numerical solution of discontinuous solution of hydrodynamic equations. Math Sbornik 47 (3), 271-306.

Han, K., Piccoli, B., Friesz, T.L., Yao, T., 2012. A continuous-time link-based kinematic wave model for dynamic traffic networks. arXiv: 1208.5141v1.

Herty, M., Klar, A., 2003. Modeling, simulation, and optimization of traffic flow networks, SIAM Journal on Scientific Computing 25 (3), 1066-1087.

Holden, H., Risebro, N.H., 1995. A mathematical model of traffic flow on a network of unidirectional roads, SIAM Journal on Mathematical Analysis 26 (4), 999-1017.

Hunt, P.B., Robertson, D.I., Bretherton, R.D., Royle, M.C., 1982. The scoot on-line traffic signal optimisation technique. Traffic Engineering and Control 23 (4), 190-192.

Improta, G., Cantarella, G.E., 1984. Control system design for an individual signalized junction. Transportation Research Part B 18 (2), 147-167.

Jin, W.-L., 2010. Continuous kinematic wave models of merging traffic flow. Transportation Research Part B 44, 1084-1103.

Jin, W.-L., Zhang, H.M., 2003. On the distribution schemes for determining flows through a merge. Transportation Research Part B 37 (6), 521-540.

Lax, P.D., 1957. Hyperbolic systems of conservation laws II. Communications on Pure and Applied Mathematics 10 (4), 537-566.

Lax, P.D., 1973. Hyperbolic systems of conservation laws and the mathematical theory of shock waves, SIAM.
Explicit formula for scalar non-linear conservation laws with boundary condition. Mathematical Models and Methods in Applied Sciences 10 (3), 265-287.

Lebacque, J., Khoshyaran, M., 1999. Modeling vehicular traffic flow on networks using macroscopic models, in Finite Volumes for Complex Applications II, 551-558, Hermes Science Publications, Paris.

Lebacque, J., Khoshyaran, M., 2002. First order macroscopic traffic flow models for networks in the context of dynamic assignment, Transportation Planning State of the Art, M. Patriksson and K. A. P. M. Labbe, eds., Kluwer Academic Publushers, Norwell, MA.

LeVeque, R.J., 1992. Numerical Methods for Conservation Laws. Birkhäuser.

Lighthill, M., Whitham, G., 1955. On kinematic waves. II. A theory of traffic flow on long crowded roads. Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences 229 (1178), 317-345.

Lin, W.H., Wang, C., 2004. An enhanced 0 - 1 mixed-integer LP formulation for traffic signal control. IEEE Transactions on Intelligent transportation systems 5 (4), 238-245.

Lo, H., 1999. A novel traffic signal control formulation. Transportation Research Part A 33, 433-448.

Lo, H., 1999. A cell-based traffic control formulation: strategies and benefits of dynamic timing plans. Transportation Science 35 (2), 148-164.

Lo, H., A dynamic traffic assignment formulation that encapsulates the cell-transmission model. Transportation and Traffic Theory, 327-350.

Moskowitz, K., 1965. Discussion of 'freeway level of service as influenced by volume and capacity characteristics' by D.R. Drew and C.J. Keese. Highway Research Record, 99, 43-44.

Nagurney, A., Qiang, Q., Nagurney, L.S., 2010. Environmental impact assessment of transportation networks with degradable links in an era of climate change. International Journal of Sustainable Transportation, 4, 154-171.

Newell, G.F., 1993. A simplified theory of kinematic waves in highway traffic, part I: General theory. Transportation Research Part B 27 (4), 281-287.

Mirchandani, P., Head, L., 2000. A real-time traffic signal control system: architecture, algorithms, and analysis. Transportation Research Part C 9, 415-432.

Penic, M. A., and J. Upchurch. TRANSYT-7F, Enhancement for Fuel Consumption, Pollution Emissions, and User Costs. In Transportation Research Record 1360, TRB, National Research Council, Washington, D.C., 1992.

Richards, P.I., 1956. Shockwaves on the highway. Operations Research 4 (1), 42-51.

Rikun, A.A. 2011. Application of robust optimization to queueing and inventory systems. Ph.D. Dissertation. Sloan School of Management. Massachusetts Institute of Technology.

Shen, W., Nie, Y., Zhang, H.M., 2007. Dynamic network simplex method for designing emergency evacuation plans. Trans. Res. Rec. No. 2022, 83-93.

Sims, A.G., Dobinson, K.W., 1980. The Sydney coordinated adaptive traffic (SCAT) system philosophy and benefits. IEEE Transactions on Vehicular Technology 29 (2), 130-137.

Szeto, W.Y., Jaber, X., Wong, S.C., 2012. Road network equilibrium approaches to environmental sustainability. Transport Reviews: A Transnational Transdisciplinary Journal 32 (4), 491-518.

Yang, H., Bell, M.G.H, 1997. Traffic restraint, road pricing and network equilibrium. Transportation Research Part B 31, 303-314.

Yang, H., Xu, W., He, B., Meng, Q., 2010. Road pricing for congestion control with unknown demand and cost functions. Transportation Research Part C 18, 157-175.

Yperman, I., S. Logghe and L. Immers, 2005. The Link Transmission Model: An Efficient Implementation of the Kinematic Wave Theory in Traffic Networks, Advanced OR and AI Methods in Transportation, Proc. 10th EWGT Meeting and 16th Mini-EURO Conference, Poznan, Poland, 122-127, Publishing House of Poznan University of Technology.